

09/869066

APPLICATION UNDER UNITED STATES PATENT LAWS

Atty. Dkt. No. PW 280384/Z 70442/UST
(M#)

531 Rec'd PCT.

21 JUN 2001

Invention: SINGLE NUCLEOTIDE POLYMORPHISM IN A PYRUVATE DEHYDROGENASE KINASE
ISOENZYME 2 (PDK2) IN HUMANS

Inventor (s): ANAND, Rakesh
MORTEN, John Edward Norris
SMITH, John Craig

Pillsbury Winthrop LLP
Intellectual Property Group
1600 Tysons Boulevard

McLean, VA 22102
Attorneys
Telephone: (703) 905-2000

This is a:

- ☐ Provisional Application
- ☐ Regular Utility Application
- ☐ Continuing Application
 - ☒ The contents of the parent are incorporated by reference
- ☒ PCT National Phase Application
- ☐ Design Application
- ☐ Reissue Application
- ☐ Plant Application
- ☐ Substitute Specification
Sub. Spec Filed _____
in App. No. _____ / _____
- ☐ Marked up Specification re
Sub. Spec. filed _____
In App. No _____ / _____

SPECIFICATION

SINGLE NUCLEOTIDE POLYMORPHISM IN A PYRUVATE DEHYDROGENASE KINASE ISOENZYME 2 (PDK2) IN HUMANS

his
AI

This invention relates to polymorphisms in the human pyruvate dehydrogenase kinase isoenzyme 2 (PDK2) gene. The invention also relates to methods and materials for analysing
5 allelic variation in the PDK2 gene, and to the use of PDK2 polymorphism in the diagnosis and treatment of diseases in which inhibition of PDK2 could be of therapeutic benefit, such as diabetes, obesity and sepsis.

Within tissues, ATP provides the energy for synthesis of complex molecules and, in muscle, for contraction. ATP is generated from the breakdown of energy-rich substrates, such
10 as glucose or long chain free fatty acids. In oxidative tissues, such as muscle, the majority of the ATP is generated from acetyl CoA which enters the citric acid cycle: thus the supply of acetyl CoA is a critical determinant of ATP production in oxidative tissues. Acetyl CoA is produced either by β -oxidation of fatty acids or as a result of glucose metabolism by the glycolytic pathway. The key regulatory enzyme in controlling the rate of acetyl CoA
15 formation from glucose is pyruvate dehydrogenase (PDH) which catalyses the oxidation of pyruvate to acetyl CoA and carbon dioxide with concomitant reduction of NAD to NADH.

PDH is an intramitochondrial multienzyme complex consisting of multiple copies of several subunits including three enzyme activities required for the completion of the conversion of pyruvate to acetyl CoA (Patel and Roche 1990; FASEB J., 4: 3224-3233). E1
20 catalyses the non-reversible removal of CO_2 from pyruvate; E2 forms acetyl CoA and E3 reduces NAD to NADH. Two additional enzyme activities are associated with the complex: a specific kinase (PDK) which is capable of phosphorylating E1 at three serine residues. Phosphorylation of a single one of the three serine residues renders the E1 inactive. In addition the complex contains a loosely-associated specific phosphatase which reverses the
25 phosphorylation: the proportion of the PDH in its active (dephosphorylated) state is therefore determined by a balance between the activity of the kinase and phosphatase. The activity of the kinase may be regulated in vivo by the relative concentrations of metabolic substrates such as NAD/NADH, CoA/acetylCoA and ADP/ATP as well as by the availability of pyruvate itself, therefore providing intimate appropriate control of substrate availability.

30 There are at least three isoenzymic forms of pyruvate dehydrogenase kinase in humans. Gudi et al (1995) J. Biol. Chem. 48, 28989-28994 reported the sequences for PDK1,

09869066-002404

- 2 -

PDK2 and PDK3. The tissue distribution of the PDK isoenzymes differed markedly in these studies. The highest levels of PDK2 mRNA were found in heart and skeletal muscle, the lowest amount in placenta and lung.

In disease states such as both non-insulin dependent (NIDDM) and insulin-dependent diabetes (IDDM), oxidation of lipids is increased with a concomitant reduction in utilisation of glucose, contributing to the hyperglycaemia. The activity of PDH is reduced in both insulin-dependent and non insulin-dependent diabetes. A further consequence of reduced PDH activity would be an increase in pyruvate concentration resulting in increased availability of lactate as a substrate for hepatic gluconeogenesis. Diabetes would be further exacerbated by impaired insulin secretion, which has been shown to be associated with reduced PDH activity in pancreatic β -cells. It is believed that increasing the activity of PDH would increase the rate of glucose oxidation and hence overall glucose utilisation, in addition to reducing hepatic glucose output.

Oxidation of glucose is capable of yielding more molecules of ATP per mole of oxygen than is oxidation of fatty acids, therefore in conditions where energy demand may exceed energy supply, such as myocardial ischaemia and reperfusion, intermittent claudication, cerebral ischaemia and reperfusion, shifting the balance of substrate utilisation in favour of glucose metabolism may be expected to improve the ability to maintain ATP levels and hence function. Activation of PDH is predicted to have this effect.

An agent which is capable of activating PDH is expected to be of benefit in treating conditions where an excess of circulating lactic acid is manifest such as in certain cases of sepsis.

The agent dichloroacetic acid which increases the activity of PDH after acute administration in animals (Vary et al., 1988; Circ. Shock, 24: 3-18) has been shown to have the predicted effects in reducing glycaemia (Stacpoole et al, 1978 N. Engl. J. Med. 298, 526-530) and as a therapy for myocardial ischaemia (Bersin and Stacpoole 1997; American Heart Journal, 134: 841-855) and lactic acidaemia (Stacpoole et al, 1983 N. Engl. J. Med 309, 390-396).

A cDNA encoding the PDK2 gene has been cloned and published by Gudi et al (1995) J. Biol. Chem. 270, 28989-28994. The sequence was submitted to the EMBL database under

EMBL Accession number: L42451 (1422 bp) and all positions herein relate to the position therein unless stated otherwise or apparent from the context.

One approach is to use knowledge of polymorphisms to help identify patients most suited to therapy with particular pharmaceutical agents (this is often termed "pharmacogenetics"). Pharmacogenetics can also be used in pharmaceutical research to assist the drug selection process. Polymorphisms are used in mapping the human genome and to elucidate the genetic component of diseases. The reader is directed to the following references for background details on pharmacogenetics and other uses of polymorphism detection: Linder *et al.* (1997), *Clinical Chemistry*, 43, 254; Marshall (1997), *Nature Biotechnology*, 15, 1249; International Patent Application WO 97/40462, Spectra Biomedical; and Schafer *et al.* (1998), *Nature Biotechnology*, 16, 33.

Clinical trials have shown that patient response to treatment with pharmaceuticals is often heterogeneous. Thus there is a need for improved approaches to pharmaceutical agent design and therapy.

The present invention is based on the discovery of a single nucleotide polymorphism (SNP) in the coding region of the human PDK2 gene and two single nucleotide polymorphisms in the 3' untranslated region (3'UTR) of the human PDK2 gene.

According to one aspect of the present invention there is provided a method for the diagnosis of a single nucleotide polymorphism in a PDK2 gene in a human, which method comprises determining the sequence of the nucleic acid of the human at one or more of positions 288, 1281 and 1357 in the PDK2 gene as defined by the positions in EMBL ACCESSION NO. L42451, and determining the status of the human by reference to polymorphism in the PDK2 gene.

The term human includes both a human having or suspected of having a PDK2-mediated disease and an asymptomatic human who may be tested for predisposition or susceptibility to such disease. At each position the human may be homozygous for an allele or the human may be a heterozygote.

The term 'PDK2-mediated disease' means any disease in which changing the level of PDK2 or changing the activity of PDK2 would be of therapeutic benefit.

The term 'PDK2 drug' means any drug which changes the level of PDK2 or changes the activity of PDK2. A drug which inhibits the activity of PDK2 is preferred.

0965906-092401
T04260-99059860

In one embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 288 is presence of C and/or T.

In another embodiment of the invention preferably the method for diagnosis described
5 herein is one in which the single nucleotide polymorphism at position 1281 is presence of G and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 1357 is presence of G and/or C.

10 The method for diagnosis is preferably one in which the sequence is determined by a method selected from amplification refractory mutation system and restriction fragment length polymorphism.

In another aspect of the invention we provide a method for the diagnosis of PDK2-mediated disease, which method comprises:

- 15 i) obtaining sample nucleic acid from an individual,
ii) detecting the presence or absence of a variant nucleotide at one or more of positions 288, 1281 and 1357 (as defined by the position in EMBL accession number L42451), in the PDK2-gene and
iii) determining the status of the individual by reference to polymorphism in the PDK2 gene.

20 Allelic variation at position 288 consists of a single base substitution from C (the published base), preferably to T. Allelic variation at position 1281 consists of a single base substitution from G (the published base), preferably to A. Allelic variation at position 1357 consists of a single base substitution from G (the published base), preferably to C. The status of the individual may be determined by reference to allelic variation at any one, two, or all
25 three positions optionally in combination with any other polymorphism in the gene that is (or becomes) known.

The test sample of nucleic acid is conveniently present in a sample of blood, bronchoalveolar lavage fluid, sputum, or other body fluid or tissue obtained from an individual. It will be appreciated that the test sample may equally comprise a nucleic acid
30 sequence corresponding to the sequence in the test sample, that is to say that all or a part of

the region in the sample nucleic acid may firstly be amplified using any convenient technique e.g. PCR, before analysis of allelic variation.

It will be apparent to the person skilled in the art that there are a large number of analytical procedures which may be used to detect the presence or absence of variant nucleotides at one or more polymorphic positions of the invention. In general, the detection of allelic variation requires a mutation discrimination technique, optionally an amplification reaction and optionally a signal generation system. Table 1 lists a number of mutation detection techniques, some based on PCR. These may be used in combination with a number of signal generation systems, a selection of which is listed in Table 2. Further amplification techniques are listed in Table 3. Many current methods for the detection of allelic variation are reviewed by Nollau *et al.*, Clin. Chem. 43, 1114-1120, 1997; and in standard textbooks, for example "Laboratory Protocols for Mutation Detection", Ed. by U. Landegren, Oxford University Press, 1996 and "PCR", 2nd Edition by Newton & Graham, BIOS Scientific Publishers Limited, 1997.

15

Abbreviations:

ALEX TM	Amplification refractory mutation system linear extension
APEX	Arrayed primer extension
ARMS TM	Amplification refractory mutation system
b-DNA	Branched DNA
CMC	Chemical mismatch cleavage
bp	base pair
COPS	Competitive oligonucleotide priming system
DGGE	Denaturing gradient gel electrophoresis
FRET	Fluorescence resonance energy transfer
IDDM	Insulin-dependent diabetes mellitus
LCR	Ligase chain reaction
MASDA	Multiple allele specific diagnostic assay
NASBA	Nucleic acid sequence based amplification
NIDDM	non-insulin dependent diabetes mellitus

OLA	Oligonucleotide ligation assay
PCR	Polymerase chain reaction
PDH	Pyruvate Dehydrogenase
PDK	Pyruvate Dehydrogenase Kinase
PDK2	Pyruvate Dehydrogenase Kinase Isoenzyme 2
PTT	Protein truncation test
RFLP	Restriction fragment length polymorphism
SDA	Strand displacement amplification
SNP	Single nucleotide polymorphism
SSCP	Single-strand conformation polymorphism analysis
SSR	Self sustained replication
TGGE	Temperature gradient gel electrophoresis
3'UTR	3' Untranslated Region

Table 1 - Mutation Detection Techniques

General: DNA sequencing, Sequencing by hybridisation

Scanning: PTT*, SSCP, DGGE, TGGE, Cleavase, Heteroduplex analysis, CMC, Enzymatic mismatch cleavage

* Note: not useful for detection of promoter polymorphisms.

Hybridisation Based:

Solid phase hybridisation: Dot blots, MASDA, Reverse dot blots, Oligonucleotide arrays (DNA Chips).

10 Solution phase hybridisation: Taqman™ - US-5210015 & US-5487972 (Hoffmann-La Roche), Molecular Beacons - Tyagi *et al* (1996), Nature Biotechnology, 14, 303; WO 95/13399 (Public Health Inst., New York).

Extension Based: ARMS™, ALEX™ - European Patent No. EP 332435 B1 (Zeneca Limited), COPS - Gibbs *et al* (1989), Nucleic Acids Research, 17, 2347.

15 **Incorporation Based:** Mini-sequencing, APEX

Restriction Enzyme Based: RFLP, Restriction site generating PCR

Ligation Based: OLA

Other: Invader assay

Table 2 - Signal Generation or Detection Systems

Fluorescence: FRET, Fluorescence quenching, Fluorescence polarisation - United Kingdom

5 Patent No. 2228998 (Zeneca Limited)

Other: Chemiluminescence, Electrochemiluminescence, Raman, Radioactivity, Colorimetric, Hybridisation protection assay, Mass spectrometry.

Table 3 - Further Amplification Methods

10 SSR, NASBA, LCR, SDA, b-DNA

Preferred mutation detection techniques include ARMSTTM, ALEXTM, COPS, Taqman, Molecular Beacons, RFLP, and restriction site based PCR and FRET techniques.

Particularly preferred methods include ARMSTTM and RFLP based methods. ARMSTTM
15 is an especially preferred method.

In a further aspect, the diagnostic methods of the invention are used to assess the efficacy of therapeutic compounds in the treatment of PDK2-mediated diseases such as diabetes, obesity, sepsis, and peripheral vascular disease.

Assays, for example reporter-based assays, may be devised to detect whether one or
20 more of the above polymorphisms affect transcription levels and/or message stability.

Individuals who carry particular allelic variants of the PDK2 gene may therefore exhibit differences in their ability to regulate protein biosynthesis under different physiological conditions and may display altered abilities to react to different diseases. In addition, differences in protein regulation arising as a result of allelic variation may have a
25 direct effect on the response of an individual to drug therapy. The diagnostic methods of the invention may be useful both to predict the clinical response to such agents and to determine therapeutic dose.

In a further aspect, the diagnostic methods of the invention, are used to assess the predisposition of an individual to diseases mediated by PDK2. This may be particularly
30 relevant in the development of diabetes, obesity, sepsis, and peripheral vascular disease and

other diseases which are mediated by PDK2. The present invention may be used to recognise individuals who are particularly at risk from developing these conditions.

Low frequency polymorphisms may be particularly useful for haplotyping as described below. A haplotype is a set of alleles found at linked polymorphic sites (such as within a gene) on a single (paternal or maternal) chromosome. If recombination within the gene is random, there may be as many as 2^n haplotypes, where 2 is the number of alleles at each SNP and n is the number of SNPs. One approach to identifying mutations or polymorphisms which are correlated with clinical response is to carry out an association study using all the haplotypes that can be identified in the population of interest. The frequency of each haplotype is limited by the frequency of its rarest allele, so that SNPs with low frequency alleles are particularly useful as markers of low frequency haplotypes. As particular mutations or polymorphisms associated with certain clinical features, such as adverse or abnormal events, are likely to be of low frequency within the population, low frequency SNPs may be particularly useful in identifying these mutations (for examples see: Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. *Ann Hum Genet* (1998) 62:481-90, De Stefano V, Dekou V, Nicaud V, Chasse JF, London J, Stansbie D, Humphries SE, and Gudnason V; and Variation at the von willebrand factor (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood* (1999) 93:4277-83, Keightley AM, Lam YM, Brady JN, Cameron CL, Lillicrap D).

In a further aspect, the diagnostic methods of the invention are used in the development of new drug therapies which selectively target one or more allelic variants of the PDK2 gene. Identification of a link between a particular allelic variant and predisposition to disease development or response to drug therapy may have a significant impact on the design of new drugs. Drugs may be designed to regulate the biological activity of variants implicated in the disease process whilst minimising effects on other variants.

In a further diagnostic aspect of the invention the presence or absence of variant nucleotides is detected by reference to the loss or gain of, optionally engineered, sites recognised by restriction enzymes. In the accompanying Example 2 we provide details of

convenient engineered restriction enzyme sites that are lost or gained as a result of a polymorphism of the invention.

According to another aspect of the present invention there is provided a nucleic acid comprising any one of the following polymorphisms:

- 5 the nucleic acid of EMBL ACCESSION No. L42451 with T at position 288 as defined by the position in EMBL ACCESSION No. L42451;
- the nucleic acid of EMBL ACCESSION No. L42451 with A at position 1281 as defined by the position in EMBL ACCESSION No. L42451;
- the nucleic acid of EMBL ACCESSION No. L42451 with C at position 1357 as defined by
- 10 the position in EMBL ACCESSION No. L42451;
- or a complementary strand thereof or an antisense sequence thereto or a fragment thereof of at least 20 bases comprising at least one polymorphism.

Fragments are at least 17 bases, more preferably at least 20 bases, more preferably at least 30 bases.

- 15 The scope of the invention does not extend to any nucleic acid as it is found in nature. A nucleic acid of the invention is preferably in isolated form, for example through being at least partially purified from any substance with which it occurs naturally (if any).

- Novel sequence disclosed herein, may be used in another embodiment of the invention to regulate expression of the gene in cells by the use of antisense constructs. To
- 20 enable methods of down-regulating expression of the gene of the present invention in mammalian cells, an example antisense expression construct can be readily constructed for instance using the pREP10 vector (Invitrogen Corporation). Transcripts are expected to inhibit translation of the gene in cells transfected with this type construct. Antisense transcripts are effective for inhibiting translation of the native gene transcript, and capable of
 - 25 inducing the effects (e.g., regulation of tissue physiology) herein described. Oligonucleotides which are complementary to and hybridizable with any portion of novel gene mRNA disclosed herein are contemplated for therapeutic use. U.S. Patent No. 5,639,595, Identification of Novel Drugs and Reagents, issued Jun. 17, 1997, wherein methods of identifying oligonucleotide sequences that display in vivo activity are thoroughly described,
 - 30 is herein incorporated by reference. Expression vectors containing random oligonucleotide sequences derived from previously known polynucleotides are transformed into cells. The

09569066-082404
TOP SECRET

cells are then assayed for a phenotype resulting from the desired activity of the oligonucleotide. Once cells with the desired phenotype have been identified, the sequence of the oligonucleotide having the desired activity can be identified. Identification may be accomplished by recovering the vector or by polymerase chain reaction (PCR) amplification and sequencing the region containing the inserted nucleic acid material. nucleotide molecules can be synthesized for antisense therapy. These antisense molecules may be DNA, stable derivatives of DNA such as phosphorothioates or methylphosphonates, RNA, stable derivatives of RNA such as 2'-O-alkylRNA, or other oligonucleotide mimetics. U.S. Patent No. 5,652,355, Hybrid Oligonucleotide Phosphorothioates, issued July 29, 1997, and U.S. Patent No. 5,652,356, Inverted Chimeric and Hybrid Oligonucleotides, issued July 29, 1997, which describe the synthesis and effect of physiologically-stable antisense molecules, are incorporated by reference. Antisense molecules may be introduced into cells by microinjection, liposome encapsulation or by expression from vectors harboring the antisense sequence.

15 The invention further provides nucleotide primers which can detect the polymorphisms of the invention.

According to another aspect of the present invention there is provided an allele specific primer capable of detecting a PDK2 gene polymorphism at one or more of positions 288, 1281 and 1357 in the PDK2 gene as defined by the positions in EMBL ACCESSION NO. L42451.

An allele specific primer is used, generally together with a constant primer, in an amplification reaction such as a PCR reaction, which provides the discrimination between alleles through selective amplification of one allele at a particular sequence position e.g. as used for ARMSTTM assays. The allele specific primer is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

An allele specific primer preferably corresponds exactly with the allele to be detected but derivatives thereof are also contemplated wherein about 6-8 of the nucleotides at the 3' terminus correspond with the allele to be detected and wherein up to 10, such as up to 8, 6, 4, 2, or 1 of the remaining nucleotides may be varied without significantly affecting the properties of the primer.

Primers may be manufactured using any convenient method of synthesis. Examples of such methods may be found in standard textbooks, for example "Protocols for Oligonucleotides and Analogues; Synthesis and Properties," Methods in Molecular Biology Series; Volume 20; Ed. Sudhir Agrawal, Humana ISBN: 0-89603-247-7; 1993; 1st Edition. If
5 required the primer(s) may be labelled to facilitate detection.

According to another aspect of the present invention there is provided an allele-specific oligonucleotide probe capable of detecting a PDK2 gene polymorphism at one or more of positions 288, 1281 and 1357 in the PDK2 gene as defined by the positions in EMBL
ACCESSION NO. L42451.

10 The allele-specific oligonucleotide probe is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

The design of such probes will be apparent to the molecular biologist of ordinary skill. Such probes are of any convenient length such as up to 50 bases, up to 40 bases, more conveniently up to 30 bases in length, such as for example 8-25 or 8-15 bases in length. In
15 general such probes will comprise base sequences entirely complementary to the corresponding wild type or variant locus in the gene. However, if required one or more mismatches may be introduced, provided that the discriminatory power of the oligonucleotide probe is not unduly affected. The probes of the invention may carry one or more labels to facilitate detection.

20 According to another aspect of the present invention there is provided a diagnostic kit comprising an allele specific oligonucleotide probe of the invention and/or an allele-specific primer of the invention.

The diagnostic kits may comprise appropriate packaging and instructions for use in the methods of the invention. Such kits may further comprise appropriate buffer(s), nucleotides,
25 and polymerase(s) such as thermostable polymerases, for example taq polymerase.

In another aspect of the invention, the single nucleotide polymorphisms of this invention may be used as genetic markers in linkage studies. This particularly applies to the polymorphisms at 288 (as defined by the position in EMBL ACCESSION NO. L42451) because of its relatively high frequency (see below).

30 According to another aspect of the present invention there is provided a method of treating a human in need of treatment with a PDK2 drug in which the method comprises:

036906-032401

- i) diagnosis of a single nucleotide polymorphism in the PDK2 gene in the human, which diagnosis comprises determining the sequence of the nucleic acid at one or more of positions 288, 1281 and 1357 of the PDK2 gene as defined by the position in EMBL accession number L42451, and determining the status of the human by reference to polymorphism in the PDK2 gene; and
- ii) administering an effective amount of a PDK2 drug.

Preferably determination of the status of the human is clinically useful. Examples of clinical usefulness include deciding which drug or drugs to administer and/or establishing the effective amount of the drug or drugs.

- 10 PDK inhibitors have been disclosed in the following publications: Whitehouse et al (1974) Biochem J. **141**, 761-774; and Espinal et al (1995) Drug Dev. Res. **35**, 130-136.

PDK inhibitors are of value in a number of disease conditions, including disease states associated with disorders of glucose utilisation such as diabetes, obesity and associated with excessive production of lactate such as encountered in sepsis and other causes of lactic acidemia. Additionally PDK inhibitors may be expected to have utility in diseases where supply of energy-rich substrates to tissues is limiting such as peripheral vascular disease, coronary failure and certain cardiac myopathies muscle ataxia, weakness.

According to another aspect of the present invention there is provided use of a PDK2 drug in the preparation of a medicament for treating a PDK2-mediated disease in a human diagnosed as having a single nucleotide polymorphism at one or more of positions 288, 1281 and 1357 in the PDK2 gene as defined by the position in EMBL accession number L42451.

According to another aspect of the present invention there is provided a pharmaceutical pack comprising PDK2 drug and instructions for administration of the drug to humans diagnostically tested for a single nucleotide polymorphism at one or more of positions 288, 1281 and 1357 in the PDK2 gene as defined by the position in EMBL accession number L42451.

According to another aspect of the present invention there is provided a computer readable medium comprising at least one novel polynucleotide sequence of the invention stored on the medium. The computer readable medium may be used, for example, in homology searching, mapping, haplotyping, genotyping or pharmacogenetic analysis or any other bioinformatic analysis. The reader is referred to Bioinformatics, A practical guide to the

analysis of genes and proteins, Edited by A D Baxevanis & B F F Ouellette, John Wiley & Sons, 1988. Any computer readable medium may be used, for example, compact disk, tape, floppy disk, hard drive or computer chips.

The polynucleotide sequences of the invention, or parts thereof, particularly those
5 relating to and identifying the single nucleotide polymorphisms identified herein represent a valuable information source, for example, to characterise individuals in terms of haplotype and other sub-groupings, such as investigation of susceptibility to treatment with particular drugs. These approaches are most easily facilitated by storing the sequence information in a computer readable medium and then using the information in standard bioinformatics
10 programs or to search sequence databases using state of the art searching tools such as "GCC". Thus, the polynucleotide sequences of the invention are particularly useful as components in databases useful for sequence identity and other search analyses. As used herein, storage of the sequence information in a computer readable medium and use in sequence databases in relation to 'polynucleotide or polynucleotide sequence of the invention'
15 covers any detectable chemical or physical characteristic of a polynucleotide of the invention that may be reduced to, converted into or stored in a tangible medium, such as a computer disk, preferably in a computer readable form. For example, chromatographic scan data or peak data, photographic scan or peak data, mass spectrographic data, sequence gel (or other) data.

20 The invention provides a computer readable medium having stored thereon one or more polynucleotide sequences of the invention. For example, a computer readable medium is provided comprising and having stored thereon a member selected from the group consisting of: a polynucleotide comprising the sequence of a polynucleotide of the invention, a polynucleotide consisting of a polynucleotide of the invention, a polynucleotide which
25 comprises part of a polynucleotide of the invention, which part includes at least one of the polymorphisms of the invention, a set of polynucleotide sequences wherein the set includes at least one polynucleotide sequence of the invention, a data set comprising or consisting of a polynucleotide sequence of the invention or a part thereof comprising at least one of the polymorphisms identified herein.

30 A computer based method is also provided for performing sequence identification, said method comprising the steps of providing a polynucleotide sequence comprising a

- 14 -

polymorphism of the invention in a computer readable medium; and comparing said polymorphism containing polynucleotide sequence to at least one other polynucleotide or polypeptide sequence to identify identity (homology), i.e. screen for the presence of a polymorphism.

5 The invention will now be illustrated but not limited by reference to the following Examples. All temperatures are in degrees Celsius.

In the Examples below, unless otherwise stated, the following methodology and materials have been applied.

AMPLITAQ™ or AMPLITAQ GOLD™ available from Perkin-Elmer Cetus, are used
10 as the source of thermostable DNA polymerase.

General molecular biology procedures can be followed from any of the methods described in "Molecular Cloning - A Laboratory Manual" Second Edition, Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory, 1989).

Electropherograms were obtained in a standard manner: data was collected by ABI377
15 data collection software and the wave form generated by ABI Prism sequencing analysis (2.1.2).

Example 1

Identification of Polymorphisms

20 1. **Methods**

c-DNA Preparation

RNA was prepared from lymphoblastoid cell lines from Caucasian donors using standard laboratory protocols (Chomczynski and Sacchi, Anal. Biochem. **162**, 156-159, 1987) and used to generate first strand cDNA (Gubler and Hoffman, Gene **25**, 263-269, 1983).

25 Template Preparation

Templates were prepared by PCR using the oligonucleotide primers and annealing temperatures set out below. The extension temperature was 72° and denaturation temperature 94°; each step was 1 minute. Generally 100 pg cDNA was used in each reaction and subjected to 40 cycles of PCR.

Fragment	Forward Oligo	Reverse Oligo	Annealing Temp	MgCl ₂
12-494	12-32	474-494	68°	2mM
380-967	380-402	945-967	67°	1mM
861-1421	861-881	1399-1421	63°	1mM

For dye-primer sequencing the forward primers were modified to include M13 forward sequence (ABI protocol P/N 402114, Applied Biosystems) at the 5' end of the oligonucleotides.

5 Dye Primer Sequencing

Dye-primer sequencing using M13 forward primer was as described in the ABI protocol P/N 402114 for the ABI Prism™ dye primer cycle sequencing core kit with "AmpliTaq FS"™ DNA polymerase, modified in that the annealing temperature was 45° and DMSO was added to the cycle sequencing mix to a final concentration of 5 %.

- 10 The extension reactions for each base were pooled, ethanol/sodium acetate precipitated, washed and resuspended in formamide loading buffer.

4.25 % Acrylamide gels were run on an automated sequencer (ABI 377, Applied Biosystems).

15 2. Results

Novel Polymorphisms

Position	Published	Variant	Amino acid change	RFLP	Frequency
288	C	T	No	eng + BsrG I or eng + Fsp I	21/58
1281	G	A	No	+Nsi I	8/44
1357	G	C	No	eng -BamHI	6/28

Frequency is the allele frequency of the variant allele in control subjects.

- 16 -

"eng" = engineered RFLP

Example 2

Engineered restriction site for detection of polymorphisms

5

Standard methodology can be used to detect the polymorphism at positions (288 and 1357 as defined by the position in EMBL ACCESSION NO. L42451) based on the materials set out below using a cDNA template.

Position	Diagnostic Fragment	Forward primer	Reverse primer
288	13-314	13-33	289-314 BsrG I or Fsp I
1357	861-1379	861-881	1358-1379 BamH I

10

Primer Sequences 5'-3'

289-314 BsrG I	acataccagctctgcaccagctgtac	SEQ ID NO.1
289-314 Fsp I	acataccagctctgcaccagctgcgc	SEQ ID NO.2
1358-1379 BamH I	cagggagaacccacccccggat	SEQ ID NO.3

15

T at position 288 creates a Bsr G I or a Fsp I site in the diagnostic fragment, 13-314, described above.

G at position 1357 creates a BamH I site in the diagnostic fragment, 861-1379, described above.

20

Example 3

Mapping of PDK2 gene to chromosome 17q21

Primer pairs in the 3' UTR and spanning intron 7 of PDK2 were used to map the gene using the G3 RH panel. The intron 7 PCR product was confirmed by sequencing. Both mapped PDK2 to 17q21 linked to marker SHGC-54600. This region of chromosome 17 is syntenic to regions associated with NIDDM and obesity in animal models (Galli J, Li LS, Glaser A, et al. Genetic analysis of non-insulin dependent diabetes mellitus in the GK rat. Nat Genet. 1996; 12:31-37; and Brockmann G, Timtchenko D, Das P, et al, Detection of QTL for body weight and body fat content in mice using genetic markers. J. Anim. Breed Genet. 1996;

09369066-082401

113: 373-379). The results provide evidence that PDK2 has a regulatory role in the control of these disease states.

Example 4

5 "Scorpions"TM assay for PDK2 SNPs 1281 and 1357

PDK2 SNPs were detected according to the "Scorpions"TM method described in Whitcombe et al., Nature Biotech 17, 804-807, 1999; with the exception that amplification conditions were 45 cycles, 94°C 45secs, 55°C 45 secs.

For SNP 1281, FAM was used for the A allele dye and Tet for the G allele. For SNP
10 1357 FAM was used for the C allele dye and Tet for the G allele.

Primer Sequences 5'-3'

SNP 1281

SEQ ID NO.4

15 Common primer ggagcccaagaacacgtccacg

SEQ ID NO.5

G allele Scorpion 5cgcggcgtcacgtaagggccgccgcg76cagtccgtcctctcaggtgcagatgtac

SEQ ID NO.6

A allele Scorpion 5cgcggcgtcacgtaagggccgccgcg76cagtccgtcctctcaggtgcagatgctt

20 SNP 1357

SEQ ID NO.7

Common primer cagtgacttccatagagacagaa

SEQ ID NO.8

C allele Scorpion 5cgcggcgaacccaccccctgctgccgcg76cctcaccatcctcctggcgc

25 SEQ ID NO.9

G allele Scorpion 5cgcggcgaacccaccccctgctgccgcg76cctcaccatcctcctggcgg

5= dye

6= hexethylene glycol

30 7= methyl red

FOI 280-99069860

Sequence Listing Free Text

	SEQ ID NO.1	<223>Description of Artificial Sequence: primer
	SEQ ID NO.2	<223>Description of Artificial Sequence: primer
	SEQ ID NO.3	<223>Description of Artificial Sequence: primer
5	SEQ ID NO.4	<223>Description of Artificial Sequence: primer
	SEQ ID NO.5	<223>Description of Artificial Sequence: primer
	SEQ ID NO.6	<223>Description of Artificial Sequence: primer
	SEQ ID NO.7	<223>Description of Artificial Sequence: primer
	SEQ ID NO.8	<223>Description of Artificial Sequence: primer
10	SEQ ID NO.9	<223>Description of Artificial Sequence: primer

09869066-082404